Abstract

The continued increases in life expectancy with no obvious deceleration, the proliferation of centenarians and appearance of supercentenarians (those aged 110 years and over) leave us in no doubt that the belief that life expectancy was limited to 85 years is now untenable. Although we may ask how long the limits to life can be pushed, the crucial question is whether the extra years gained year on year in life expectancy are healthy years. This paper begins by reviewing what was historically believed to be the theoretical relationships between life expectancy and healthy life expectancy. We debate how current knowledge of mortality rates in the old and oldest old, the trends in healthy life expectancy, and the gap between the genders shed light upon these theoretical models, discussing the fact that different models may exist in different cohorts of the same population. The paper closes with some speculations on how we might monitor the evolution of healthy life expectancy more closely, particularly in those countries still early on in the ageing transition.

Background

The debate about the relationship between life expectancy and healthy life expectancy and the move from simply considering the length of life to the quality of remaining life began over two decades ago, though the debate at that time was predominantly theoretical with little data to prove or disprove theories. As mortality rates at older ages were observed to be falling, three main theories emerged. Gruenberg (1977) predicted a pandemic of chronic diseases, or expansion of morbidity, through the progress in medical care extending the life of those with disease and disability as well as life expectancy increasing to the ages at which disease and disability are more likely. Fries (1980) on the other hand gave a more optimistic outlook, that through behaviour change and prevention, we could postpone the onset of disease and disability closer to the end of life: the compression of morbidity scenario. Central to this argument was the existence of the limit to life expectancy, then thought to be around 85 years. Manton (1982) proposed an intermediate view between these extremes of a dynamic equilibrium, where the prevalence of disability may increase as mortality falls, but the severity of disability is reduced.

The dynamic equilibrium scenario points much more clearly to the crucial question of whether the extra years of life are good quality years or not and for the need to introduce a notion of severity. Disability is, at the same time, an indicator of the severity of morbid states and an indicator of the quality of years lived. The introduction of disability into models of health significantly clarified the relationships between disease and mortality leading to the partitioning of years of life into those with and without disability and to the development of health expectancy as a population health indicator to elucidate the three scenarios. This can be seen from Figure 1 (WHO, 2004) which distinguishes three elements: total survival, disability-free survival and survival without disabling chronic disease, leading to life expectancy (LE: area under the ‘mortality’ curve), disability-free life expectancy (DFLE: area under the ‘disability’ curve) and life expectancy without chronic disease (DisFLE: area under the ‘morbidity’ curve). The difference between LE and DFLE measures life expectancy with a disability, the difference between LE and DisFLE measures life expectancy with at least one chronic disease and the difference between DisFLE and DFLE measures life expectancy with at least one chronic disease but without disability.

Figure 1. General model of health transitions [WHO, 1984] The observed mortality and hypothetical morbidity and disability survival curves for women in the United States of America in 1980.
Health expectancies, of which disability-free life expectancy (DFLE) is one, thus provide a means of dividing life expectancy into life spent in various states of good and bad health. Since a health expectancy combines life expectancy with a health concept, there are as many possible health expectancies as health concepts. So it is possible to construct a number of indicators that might measure healthy life including: life expectancy ‘in good perceived health’; ‘disease-free’ life expectancy, of which dementia-free life expectancy is perhaps the most well-known (Ritchie, 1991); ‘active life expectancy’, without loss in the performance for daily living activities (ADLs, Katz et al, 1983); and life expectancy ‘without significant risk factor damage’ (Manton, 1989) or more simply without ‘risk’ (Rowe, 1990) to assess what has been called ‘successful ageing’ (Rowe & Kahn, 1997). Additional models of the ageing process enable the introduction of concepts such as survival without loss of autonomy (Grimley Evans, 1983). Furthermore, several levels of severity can be introduced for a single concept as illustrated above for disability.

Such indicators represent the increasing focus on indicators of the quality of life lived (life spent in a healthy state) rather than, as previously, on the quantity (life expectancy). Sanders (1964) proposed the notion of health expectancy and the first method of calculation followed in 1971 (Sullivan, 1971). Since that time, health expectancies have been used increasingly in industrialized countries to assess the evolution of the populations’ health status, in particular that of older people. Being independent of the size of populations and of their age structure, health expectancies provide a direct comparison of the different groups that make up populations: e.g. sexes, socio-professional categories, regions. Most importantly in the context of this paper, they are a useful tool to address whether or not the gains in life expectancy are being accompanied with an increase in time lived in bad health through the three scenarios.

The compression of morbidity or disability scenario as defined by Fries (1980) corresponded to the survival curve (Figure 1) taking on a more rectangular shape as the maximum survival remained the same but the initial decline in survival occurred at a progressively later age. Following this change in life expectancy, for a compression of morbidity, the survival curves without disease and without disability have to move much closer to the total survival curve. It has subsequently been established that the rectangularization of the survival curve (and therefore a limit to life expectancy), is not necessary for compression of disability, but simply that LE is increasing less rapidly than DFLE (Palmore, 1987; Fries et al, 2004). The expansion of morbidity scenario corresponds to the survival curve moving to the right when the disability and morbidity curves move little, thus LE increases more rapidly than DisFLE or DFLE. Manton’s (1982) theory of dynamic equilibrium is an illustration of the case of France in the 1980s, when LE and DFLE increased by the same amount but DisFLE remained constant (Robine et al, 1996).

Although the prevalence of chronic disease increased, the increase was due to less severe disease with consequently less disability.

A good target for public health policy in the light of our ageing populations would therefore seem to be to compensate for the gain in life expectancy by an equivalent increase in healthy life expectancy. In this case all the gain is healthy, changing the proportion of good and bad years and leading to a relative compression of morbidity (Robine & Mathers, 1993). But is this enough? Since in this case the number of years spent in bad health would remain the same, there would be no absolute compression of morbidity. Thus the target should be more than to simply substitute the increase in life expectancy by the same increase in healthy life expectancy. We will discuss alongside current evidence, two ways in which this might happen: firstly through a limit to life expectancy and secondly through healthy life expectancy increasing faster than life expectancy but with no limit to life expectancy.

Limits to life expectancy

Although compression of mortality is not a prerequisite for compression of morbidity, a deceleration of the growth in life expectancy will help to accentuate compression of morbidity (Fries et al, 2004). One reason why the gains in life expectancy may slow down is that a limit to life expectancy exists and that we are getting closer to it. During the early 1980s and for some time later, the limit was thought to be 85 years (Hayflick, 1981; Olshansky et al, 1990; Fries et al, 2004) though this has recently been reached by Japanese women (Robine et al, 2003b). Life expectancy appears to be increasing at the rate of about three months each year in those countries who are most advanced in the ageing transition. If these increases are extrapolated the life expectancy of women will reach 100 in at least one country by 2060 (Oeppen & Vaupel, 2002). Though there is still much debate around this (Leslie, 2003), there appears to be no deceleration of the year-on-year gains in life expectancy in those countries with the highest life expectancies: Japan, Hong Kong, France and Switzerland (Hong Kong CSD, 2002; Heiniger & Wanders, 2002; MHW Japan, 2003; Beaumel et al, 2004). The exponential increase in centenarian numbers noted in many countries (Robine & Caselli, 2005), the growth in the number of supercentenarians, those reaching the age of 110 years, and the fall in mortality at the oldest ages (Robine et al, 2003b) lends further evidence to the fact that a limit to human longevity has not yet been reached.

Since Gompertz (1825), the traditional view has been that mortality increases exponentially with age, mortality rates doubling approximately every eight years. In 1950, from extrapolation of the then current data, Vincent showed that 110 years was the age by which the probability of dying within a year would tend to 100% (Vincent, 1951). Of course this theory was put forward at a time of few oldest
Healthy life expectancy increasing faster than life expectancy

When the scenarios of the relationships between life expectancy and healthy life expectancy were developed there was little reliable comparative data on disability prevalence. Trends in disability are now available for many countries, but a major problem for synthesizing the evidence is the variety of definitions of disability in use worldwide. In addition, there are far more reports on trends in disability prevalence per se rather than incorporating these into estimates of DFLE. In our synthesis we will concentrate on trends in DFLE as it is difficult to assess whether observed trends in disability prevalence will produce DFLEs that are increasing faster than LE.

The most comprehensive review of DFLE worldwide, which attempted to harmonise the definitions of disability, was published as part of the ten years of research of REVES, the International Network of Health Expectancy and the Disability Process (Robine et al, 2003a). Countries with series of DFLEs covering parts of the period between 1968 and 1995 are USA, UK, Finland, Australia, France, New Zealand, Germany and Canada, whilst USA, UK, Australia, France, Canada could also differentiate severe disability-free life along with an extra series at this level for Japan. These series show (i) clear increasing trends in life expectancy at age 65 years for both men and women for all the countries with relatively little variation between them; (ii) on the whole a relatively constant trend in DFLE at age 65 years for both men and women, though Australia demonstrated a decline in DFLE; (iii) clear increasing trends for severe DFLE in line with LE increases. Thus these data point consistently to the dynamic equilibrium scenario over the last few decades. Switzerland and the Netherlands also have series of severe disability prevalence and these show reductions in severe disability in Switzerland and little change in the Netherlands, both consistent with dynamic equilibrium (Robine et al, 2003c). New series for Denmark (Bronnum-Hansen, 2005) have less clear increases in life expectancy and larger DFLE increase supporting a compression of morbidity, a good illustration of how compression may occur if life expectancy increases slow down. There is also some diversity within countries in which series exist from multiple sources, for example, the US, where compression of disability has also been demonstrated (Manton & Lamb, 2005). Deeg has suggested that such diversity may depend on the initial levels of disability, as in the US initial disability levels were high providing more potential for reduction and therefore compression, whilst in Australia and the Netherlands, disability levels were low with less potential for reduction and more sensitivity to local period influences (Deeg et al, 2004).

As stated earlier, a considerable problem with synthesizing evidence on DFLE across countries is the myriad of definitions of disability used. Within Europe the European Household Community Panel was used to compare DFLE
between 1995 and 2003 in the then 14 countries comprising the European Union (Luxembourg excepted). The same item on disability ‘Are you hampered in your daily activities by any physical or mental health problem, illness or disability?’ was asked of all subjects in all countries and the same method of calculating LE was used, though the estimates of DFLE may be subject to other biases including non-optimal translation from the original English, differential attrition and lack of institutional population. The latter however will be less of an issue when looking at trends as substantial sudden changes in the proportion of those in institutions or their health are unlikely. As the earlier REVES data show worldwide, LE across the EU showed clear increases over time with less variability between countries than DFLE. When the proportion of life spent free of disability was considered, countries fell into three groups, those where values over the period had: increased 5% or more (men: Austria, Belgium, Finland, Germany, Italy, Spain; women: Italy, Sweden); decreased by 5% or more (men: Denmark, Sweden, UK; women: Germany, Greece, Ireland, Netherlands, Portugal); increased or decreased by less than 5% (men: France, Greece, Ireland, Netherlands, Portugal; women: Austria, Belgium, Denmark, Finland, France, Spain, UK) (Figure 3). What is interesting is that there was little overlap in these groups for men and women suggesting gender differences in the relationship between life expectancy and healthy life expectancy. We shall return to this debate later in the next section.

Of course disability is only one facet of health. Few series are available on other health dimensions, the exception being self-perceived health where series are available for the UK (ONS, 2005), Netherlands (Perenboom, 2004) and Austria (Doblhammer & Kytir, 2001). There is again a suggestion that where life expectancy is low there is more potential for improving healthy life expectancy, as in the case of Austria. We shall return to this debate later.

Although the evidence reviewed thus far shows considerable variation both between countries and between men and women in the relationship between life expectancy and healthy life expectancy, the greatest part of it seems to be demonstrating dynamic equilibrium, with an increase in disability at milder levels. However, the question we wished to address in this section was whether healthy life expectancy was likely to increase faster than life expectancy in order to attain compression of morbidity. If disability levels are initially high and/or life expectancy is increasing very slowly, there appears to be more likelihood of compression but the question still remains as to how disability levels can be reduced. Most models of the disablement process place diseases at the beginning of the process. Thus, in addition to the social and environmental changes that might take place to alleviate disability, since disability is also a social construct, an obvious target for disability reduction is those diseases that infer the greatest disability.

Many studies have modelled the effect of diseases and conditions on mortality in older people but longitudinal data sets that can model the contribution of diseases to disability have been rare, not least because the low prevalence of some conditions means they can only be studied with large sample sizes. Though formal meta-analytic techniques were impossible due to the variety of definitions of disability used, Stuck et al (1999) undertook the most comprehensive review of studies elucidating the contributions of diseases to disability in older people, though the studies often concentrated on single diseases. This review was later updated by Spiers et al (2005) with, in addition, modelling of the UK MRC Cognitive Function and Ageing Study to assess the relative contributions of 11 different diseases and conditions to both disability and mortality.

There have been a number of projections of the number of older people with disability (Howse, 2005) but these have generally (i) considered only single diseases with an emphasis on dementia and (ii) not explicitly calculated...
trends in DFLE but have tended to infer trends in DFLE from separate life expectancy and disability trends. In contrast Wagener et al (2001) simulate the impact of reducing mortality and/or activity limitation rates on healthy life expectancy in the US. The paper greatly clarifies where potential public health targets should lie and provides strong evidence for shifting the focus from mortality to morbidity declines. One year free of activity limitation could be gained from a 12% reduction in mortality at all ages, a 7% reduction in morbidity or a 5% reduction in both. Reductions later in life had greater potential to impact in healthy life expectancy since it is precisely this age group where morbidity is highest. Furthermore it is reductions in mortality at older ages that impacts detrimentally on healthy life expectancy as years saved at older ages tend to be lived in poor health. The impact of specific diseases is however only inferred.

Over the last 20 years, there have been a series of studies assessing the potential gains in life expectancy through the elimination of diseases (Colvez & Blanchet, 1983; Bone et al, 1995; Nusselder et al, 1996; Murray & Lopez, 1997; Mathers, 1999). However since all these studies were based on mortality ascertainted from death certificates, the chronic disabling conditions were under-represented compared to fatal conditions. What is obvious is that conditions such as osteoarthritis whose prevalence increases with age will make increases in DFLE difficult to attain with further population ageing, since osteoarthritis is one of the commonest causes of disability in older people with little impact on mortality.

Gender gap
One way in which we might speculate how life and health expectancy may evolve is to explore the difference in these quantities between men and women. Increase in life expectancy at birth through the 20th century was accompanied by the emergence of a gap in favour of women which progressively widened to about 7 years: the longevity gender gap. Health expectancy calculations, paradoxically, do not confirm this gap in terms of morbidity or disability as studies consistently show that a greater percentage of women are disabled compared to men. Although, it is not clear whether there are gender differences in the reporting of disability, the prevalence of total disability in older women can be estimated to be approximately 50 percent higher than in men. This excess of morbidity and disability in women reduces or cancels their longevity advantage when assessing the quality of the years lived: the disability gap. Many health researchers simply concluded that women live longer but in poorer health than men (Barer, 1994; Franceschi et al, 2000), in line with a trade-off between quantity and quality of life. The female survival advantage can be tracked back to conception with a very large ratio of males per female leading to the actual ratio of 105 males for 100 females at birth (Moore & Persaud, 1998). For the time being there is no satisfactory explanation either for the longevity gap or for its paradoxical opposite disability gap.

In Switzerland, a country having one of the highest life expectancies at birth, the longevity gap in favour of women was weak at around 3 years and varied little before the 1930s from when it widened to reach a maximum of 7 years by 1990. However, during the last decade of the 20th century, it fell back to 5.6 years (Robine & Paccard, 2005). Similarly in Japan, a gap of 7 years was reached in 2000 and since then has plateaued at that level (Robine et al, 2003b). Although the US demonstrated a similar small gender gap in longevity before the 1930s, the gap widened to reach a maximum of about 7 years as early as in 1970s. This now relatively constant gap could be explained by the fact that female life expectancy, so high today, may be near its maximum.

The widening of the human gender gap is a modern phenomenon probably due to 20th century social and economic transformation, apparently benefiting women more than men. It also seems to correspond to a transitional period between conditions when the greatest danger to life was starvation and infectious diseases to modern conditions when the greatest danger is opulence and overnutrition. The question then is why women can derive a greater or a faster benefit from these changes in living conditions. Another question is the size of the residual gap in the future when men will possibly have caught up (Miller, 1986; Smith, 1989; Newman & Brach, 2001).

A number of biological and social factors have been proposed to explain gender-specific behaviours, habits and beliefs which could lead to the observed longevity gaps. Most researchers focus only on the longevity gap and favour a simple biological or social theory such as the chromosomal, hormonal, oxidative and replicative theories on the biological side, and stress-related job and gender roles, smoking and risk behaviour, social constructionist and feminist theories on the social science side. (Boucot, 1965; Miller, 1986; Wallace, 1996; Courtenay, 2000). Only a few look for a coherent explanation of both the longevity and the disability gaps.

The basic biological variables and factors have not changed significantly through the 20th century, and therefore cannot explain the widening of the gaps (Miller, 1986). The longevity gap could be a result of the difference in body size between men and women, needing a different number of replications and then leading to a different length of telomeres. Thus change in body size through the 20th century would explain the widening of the gaps, men being closer to their replicative limits (Stindl, 2004). Several authors underline diet as a major change and notice increasing weight gain in both men and women in recent years (Newman & Brach, 2001).

On the social science side, smoking has, for many years, been the main explanation for the longevity gap. The
current cohorts had very different smoking history related to gender and social status. Smoking became popular for men through WWI but remained expensive until WWII. Later still, around the 1960s or 1970s, many educated people stopped or reduced smoking when rates in lower social classes and women began to rise. If trends persist, the longevity gap should reduce as women approach the same total smoking years as men (Miller, 1986). Notwithstanding, men and women engage in different social practices to demonstrate their respective masculinity and femininity. Men’s beliefs and behaviours tend to undermine their health and longevity: smoking, heavy drinking, fast driving, overweight or delaying medical contact in case of health problems. Women’s beliefs and practice, on the other hand, tend to reinforce their health (health-promoting behaviours and healthy lifestyle patterns): using safety belts, healthy exercising habits, health screenings, awareness of medical conditions, use of vitamin and mineral supplements (Courtenay, 2000).

With regard to the disability gap, explanations have included reporting bias, higher rates of disabling diseases (including depression and dementia, arthritis), possible physiologic differences (including loss of muscle mass and sarcopenia, reduced mobility, increased risk of fall and increased risk to be injured when they fall and to get a hip fracture, increased body fat) and behavioural factors (women having less social support and assistance), presence of more comorbidities or chronic health problems. But at every level of comorbidity women have greater disability (Newman & Brach, 2001).

Eventually the survival advantage in women results in a larger number of women living through the different ages of life including old age. On the one hand, at any level of functioning women survive better but on the other hand, living longer, they accumulate more comorbid conditions. These mechanisms, better survival and longer life, could explain the higher rates of disability among women. However, it is possible that for any morbidity level, men are more active, displaying less disability as well as, perhaps, engaging in risk behaviours that shorten their life.

**Health transition**

Is it possible to arrange all these pieces of information in a harmonious picture? A number of authors support the idea of trade-off between quantity and quality, especially in economics. But is it possible to apply such a concept to the quantity and quality of life? Studies of LE and DFLE by socio-economic status have repeatedly shown that quantity and quality of life go together, those having the highest SES having the longest LE, DFLE and also life expectancy with disability. Although, in general, those with the highest SES have also the highest proportion of years free of disability within LE.

The demographic and epidemiologic theories of population health transition may offer an alternative explanation to the apparent contradictory results if we accept the idea that countries, genders and socio-economic groups within countries may be at different stages of a general health transition and that this transition is a permanent dynamic process of population changes over time. Thus, it has been recently proposed that the lengthening of adult life durations may be accompanied by a kind of circling back, where, first, sick people survive into older ages and disability rises, then the health status of older people is improved through various means and the number of years lived with disability decreases, but finally, the number of years lived with disability rises again when the average age of death rises to the extent that many people spend their last years at advanced old age burdened with multiple chronic diseases and frailty (Guralnik et al., 2004; Robine & Michel, 2004).

**Future**

Whatever is the correct explanation for the current patterns, the ongoing accumulation of extremely old persons will force us to change our conception of health and associated notions such as risk factors, morbidity, causes of death, and recovery. For instance it is loss of weight, not gain, that is one of the most serious risk factors for the very old.

Monitoring compression of morbidity or more simply the quality of the years lived at always older ages will not only need new data but also harmonized data if we want to be sure that future observed differences and contradictory results between countries are not due to difference in design. The International Association of Gerontology (IAG) is in the process of launching a World Ageing Survey (WAS) in collaboration with the United Nations to provide such a unified design. The goal of the WAS will be to monitor ageing at the global level through a sustainable system of cross-sectional surveys, repeated every five years in a representative sample of countries and covering the major dimensions of ageing and well-being. By this means WAS will be a global surveillance system for healthy active life at older ages providing the harmonized data that we need to definitively clarify the relationship between increase in life expectancy and healthy life expectancy.


**References**


Correspondence:
Dr Jean-Marie Robine,
Val d’Aurelle, Parc Euromedicine, 34298, Montpellier, cedex5, France.
Email: jmrobine@valdorel.fnclcc.fr